LETTER TO THE EDITOR

Testing the Phylogenetic Stability of Early Tetrapods

Clack (1998) described *Eucritta melanolimnetes*, an Early Carboniferous tetrapod, and used parsimony to investigate its evolutionary relationships to other early tetrapods. Through comparison of a most parsimonious tree and nine trees of one additional step in length, Clack concluded that “The phylogeny from the current data set is not particularly robust” and claimed that “This instability undoubtedly results from the extreme degree of character conflict that *Eucritta* presents.” We have developed measures of leaf and overall phylogenetic stability that allow such claims about the impact of individual leaves to be tested.

Phylogenetic trees represent hypotheses of evolutionary relationships for a set of terminal taxa or leaves. Most studies of the robustness of phylogenetic trees focus on the support for, or stability of, clades. However, trees can also be thought of as collections of less inclusive hypotheses of phylogenetic relationships, such as 3-taxon statements, with implications for the assessment of support (Wilkinson, 1994). For example, a single unstable leaf can result in clades with minimum robustness despite strong support for the phylogenetic relationships of the remaining leaves (Wilkinson, 1996). Measures of leaf stability allow the contribution of each leaf to overall phylogenetic stability to be determined.

The rationale for the measures is very simple. The phylogenetic relationships of a set of leaves are a function of the relationships among each subset of three leaves (triplet). The support for the relationships within each triplet provides a measure of the stability of that triplet. Stable and unstable leaves will tend to occur in stable and unstable triplets, respectively. The average stability of the triplets including a leaf provides a measure of the stability of that leaf. Similarly, the average stability of all triplets is a measure of overall phylogenetic stability. Leaf and phylogenetic stability measures can utilize any numerical technique for assessing the support for phylogenetic relationships. Here, we use a measure based on the widely used technique of bootstrapping (Felsenstein, 1985). The bootstrap proportion (BP) of any phylogenetic hypothesis is the proportion of bootstrap trees that include the hypothesis (or a weighted proportion in the case of multiple trees from a single bootstrap replicate). There are three possible 3-taxon statements for each triplet. The measure of triplet stability used here is the absolute difference between the BPs of the two best-supported 3-taxon statements for the triplet.

We reanalysed the data from Clack (1998) with uninformative characters removed. Bootstrapping was performed using PAUP 3.1.1 (Swofford, 1993) with 500 replicates, 10 random addition sequences and TBR branch swapping. Leaf and phylogenetic stabilities were determined using RadCon (Thorley and Page, 1999). Two leaves, *Crassigyrinus* and *Whatcheeria*, are less stable than *Eucritta* (Fig. 1). Interestingly, earlier authors, albeit in reference to earlier reconstructions, described these leaves as “uniquely primitive and aberrant” (Panchen, 1985) and of “uncertain phylogenetic affinities” (Lombard and Bolt, 1995).

Although *Crassigyrinus* and *Whatcheeria* are less stable than *Eucritta*, it is possible that their instability is due to character conflict introduced by *Eucritta*. In order to test the impact of the character data for individual leaves on overall phylogenetic stability we compared the stabilities of all triplets when the leaf is not included in the analysis with their stabilities when the leaf is included. Triplet stability measures when the leaf
FIG. 1. Bootstrap consensus tree of early tetrapods with clade bootstrap proportions and leaf stabilities.

is included were determined by pruning the leaf from the set of bootstrap trees and condensing identical trees from the same bootstrap replicate. The weight of each remaining tree was then set to be the inverse of the number of non-identical trees in the replicate. Statistical significance was tested using the Wilcoxon signed ranks tests (two-tailed) of the differences in stability of the triplets when the leaf is included or not included. Removing Eucritta or Whatcheeria increased overall phylogenetic stability but the differences were not statistically significant (increase = 0.20%, \( p = 0.47 \) and 0.58%, \( p = 0.77 \), respectively). In contrast, removal of Crassigyrinus significantly increased stability (1.88%, \( p = 0.003 \)).

Assessing the strengths and weaknesses of inferred relationships is an important element of phylogenetics. Leaf and phylogenetic stability measures address aspects of hypothesis quality (with respect to given data) that are not revealed by measures of clade support and, thus contribute to a more comprehensive understanding of phylogenetic hypotheses. For example, they can be used to identify unstable and potentially problematic leaves and they enhance our capacity to test claims regarding the influence of specific leaves. In this case, our results demonstrate that Crassigyrinus and Whatcheeria are the most unstable leaves in Clack’s (1998) phylogeny and that its overall instability is neither solely nor primarily due to Eucritta.

Joseph L. Thorley
School of Biological Sciences, University of Bristol, Bristol, BS8 1UG, U.K. and Department of Zoology, The Natural History Museum, London, SW7 5BD, U.K.

Mark Wilkinson
Department of Zoology, The Natural History Museum, London, SW7 5BD, U.K.

(Received on 12 May 1999, Accepted in revised form on 13 July 1999)

REFERENCES


